

# What is the process for calculating effect size in Stata for regression?

Authored by  
**stats writer**

July 1, 2024

## RECOMMENDED CITATION

stats writer (2024). *What is the process for calculating effect size in Stata for regression?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=163618>

The process for calculating effect size in Stata for regression involves using the standardized coefficient (beta) from the regression output and the standard deviation of the dependent variable. This is done by dividing the beta by the standard deviation, which results in the effect size (also known as Cohen's d). This value represents the magnitude of the relationship between the independent and dependent variables and can be compared across different studies. Stata also provides a feature for calculating other effect size measures, such as R-squared and partial eta squared, to further examine the strength of the relationship. Overall, the process for calculating effect size in Stata for regression allows for a standardized and quantitative measure of the impact of an independent variable on a dependent variable.

## How can I compute effect size in Stata for regression? | Stata FAQ

Two of the more common measures of effect size for regression analysis are eta<sup>2</sup> and partial eta<sup>2</sup>. Eta<sup>2</sup> is the proportion of the total variance that is attributed to an effect or set of effects. Partial eta<sup>2</sup> is the proportion of effect + error variance that is attributable to the effect. The formula differs from the eta squared formula in that the denominator includes the SSEffect plus the SSEError rather than the SSTotal. The Stata regress postestimation command `estat esize` can be used to estimate eta<sup>2</sup> for the model and partial eta<sup>2</sup> for each effect in the model.

Below, we run a linear regression analysis the hsbdemo dataset.

use <https://stats.idre.ucla.edu/stat/data/hsbdemo>, clear

regress write i.female read math i.prog

Source | SS df MS Number of obs = 200

-----+----- F( 5, 194) = 45.01

Model | 9602.28627 5 1920.45725 Prob > F = 0.0000

Residual | 8276.58873 194 42.6628285 R-squared =  
0.5371

-----+----- Adj R-squared = 0.5251

Total | 17878.875 199 89.843593 Root MSE = 6.5317

-----+-----  
write | Coef. Std. Err. t P>|t|

-----+-----  
1.female | 5.384982 .929572 5.79 0.000 3.551617  
7.218346

read | .3069424 .0611262 5.02 0.000 .1863852 .4274996

math | .3603705 .0690064 5.22 0.000 .2242715 .4964695

|  
prog |

2 | .436372 1.230379 0.35 0.723 -1.990265 2.863009

3 | -2.219748 1.359353 -1.63 0.104 -4.900756 .4612603

|  
\_cons | 15.16272 3.225088 4.70 0.000 8.801985 21.52346

We follow the regress command with `estat esize`, which displays estimates and confidence intervals for eta2 for the model and partial eta2 for each effect in the model.

`estat esize`

Effect sizes for linear models

Source	Eta-Squared	df
Model	.5370744	5
female	.1474719	1
read	.115024	1
math	.1232518	1
prog	.0232415	2

	.433662	.6003297
--	---------	----------

	.0667184	.2382202
--	----------	----------

	.043701	.2017348
--	---------	----------

	.0493088	.2111699
--	----------	----------

	.0732192
--	----------

An `anova` table of this regression allows us to see how eta2 and partial eta2 are calculated.

`anova write i.female c.read c.math i.prog`

## Effect sizes for linear models

Number of obs = 200 R-squared = 0.5371

Root MSE = 6.53168 Adj R-squared = 0.5251

Source | Partial SS df MS F Prob>F

```
-----+-----
Model | 9602.2863 5 1920.4573 45.01 0.0000
|
female | 1431.7 1 1431.7 33.56 0.0000
read | 1075.743 1 1075.743 25.21 0.0000
math | 1163.5091 1 1163.5091 27.27 0.0000
prog | 196.93763 2 98.468814 2.31 0.1022
|
Residual | 8276.5887 194 42.662829
-----+-----
Total | 17878.875 199 89.843593
```

The model  $\eta^2$  is  $SS_{Model}/SS_{total} = 9602.2863/17878.875 = 0.53707441$ . This matches the estimated for R-squared. The familiar interpretation is that the model explains 53.71% of the total variance of write.

Each partial  $\eta^2$  is  $SS_{Effect}/(SS_{Effect}+SS_{Error})$ . The

**SSError for all of these terms is SSResidual. We can thus calculate partial eta<sup>2</sup> for female =  $\text{SSEffect}/(\text{SSEffect}+\text{SSError}) = 1431.7/(1431.7+8276.5887) = 0.14747192$ . We can interpret this to mean that about 14.75% of the variance unexplained by effects other than female is explained by the female effect. If we need estimates of eta<sup>2</sup> for each effect, it is simply  $\text{SSEffect}/\text{SSTotal}$ .**

ARABPSYCHOLOGY.COM